EFFICIENT DOCUMENT BOUNDARY DETERMINATION

BACKGROUND OF THE INVENTION

The present application relates to document boundary determination.

5   Optical scanners operate by imaging an object, typically in the form of a sheet of paper, document, or other form of medium with a light source. The optical scanner senses a resultant light signal from the medium with an optical sensor array that includes pixel elements generating a data signal representative of the intensity of light impinging thereon for a corresponding portion of the imaged object. The data signals

10   from the array are then processed (typically digitized) and utilized by a utilization apparatus or stored on a suitable medium such as a hard drive of a computer system for subsequent display and/or manipulation.

Various types of photo sensor devices may be used in optical scanners. For example, a commonly used photo sensor device is the charge coupled device (CCD),

15   which builds up an electrical charge in response to exposure to light. The size of the electrical charge built up is dependent on the intensity and duration of the light exposure. In optical scanners, CCD cells are typically arranged in linear arrays. Each cell or "pixel" has a portion of a scan line image impinged thereon as the scan line sweeps across the scanned object. The charge build up in each of the pixels is measured and discharged at

20   regular "sampling" intervals.

The image of a scan line portion of a document is projected onto the scanner's linear sensor array by scanner optics. In CCD scanners, the scanner optics

5 typically comprise an imaging lens which typically reduces the size of the projected

image from the original size of the document. Pixels in a scanner linear photo sensor

array are aligned in a direction perpendicular to the "scan" direction, i.e. the paper or

scanner movement direction for scanning of the image.

At any instant when an object is being scanned, each pixel in the sensor

10 array has a corresponding area on the object which is being imaged thereon. This

corresponding area on the scanned object is referred to as an "object pixel." An area on a

scanned object corresponding in area to the entire area of the linear sensor array is

referred to as an "object scan line" or "scan line." For descriptive purposes, a scanned

object is considered to have a series of fixed adjacently positioned scan lines. Scanners

15 are typically operated at a scan line sweep rate such that one scan line width is traversed

during each sampling interval.

Some optical scanning machines include an automatic document feeder for

feeding a document past the optical array. Other optical scanners machines are known as

"flat-bed" scanners, wherein a document is placed on a fixed platen for scanning, which

20 occurs by moving the sensor array relative to the fixed document.

It is advantageous in various applications to sense the location of a

document edge. In a printer, for example, the print area differs depending on whether the

printing is on envelopes, name card paper, letter sized paper, and so on. The prediction of

the print area assists in driving the print head. The print area can be identified by sensing

25 the media edges. By identifying the document area, proper clipping can be made on both

sides when printing. In a scanner, detection of the document edges can assist by placing

5   the image area properly on the page, and by reducing the scan memory size by clipping

the empty regions. Also, by detecting the edge position in the direction of document

movement, the document skew can be estimated and used to redirect the scanned image

in print. This will produce a more pleasing output from the scanning process. In a

copier, sensing the size of a document permits scaling of the input document to the

10   maximum size that will fit on an output page. In addition, multi-function machines

combine in a single machine the functions of printing and optical scanning with

automatic document/sheet feeders.

If a document is misaligned with respect to the optical sensor, the resultant

image is similarly skewed. Because the contents of a document page are usually aligned

15   with the page itself, a skewed page usually results in a misalignment with the optical

sensor.

Pasco et al., U.S. Patent No. 5,818,976, disclose a system for skew and

size/shape detection of a document. The system performs the following basic steps, (1)

detects points near the edge of the page image, (2) fits lines to establish a closed contour,

20   and (3) defines a polygon with sides coincident with the lines of the closed contour. The

polygon defines the size and shape of the page image. With respect to detecting the

edges of the page the system uses a background (platen backing cover) that contrasts well

with the page, e.g., a black (or gray or patterned) background and white documents. Then

the system analyzes the image to determine the edges of the image. Unfortunately, this

25   requires specialized hardware to determine the edge of the image and thus is unsuitable

for general purpose scanning devices. If a contrasting background is not used, Pasco et

5   al. suggest the use of electro-mechanical switches or optical switches arranged to sense

the location of edges of each page in conjunction with scanning. Likewise, this requires

specialized hardware to determine the edge of the image and thus is unsuitable for general

purpose scanning devices.

What is desired, therefore, is a system that can determine the general

10   bounding region of a document without additional specialized hardware.

DETAILED DESCRIPTION OF THE DRAWINGS

FIG. 1 is an exemplary illustration of a scanner, document and cover.

FIGS. 2A, 2B AND 2C illustrate sample images.

15   FIG. 3 is an exemplary process flow chart.

FIGS. 4A and 4B illustrate stat buffers.

FIGS. 5A and 5B illustrate accumulation buffers.

FIGS. 6A and 6B illustrate smoothed buffers.

20   DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

The present inventors considered the existing prior art systems, which

generally use specialized physical devices such as sheet feeders, document delivery

systems, specially designed platen covers, multiple light sources, etc. Each of the

systems is unsuitable for general purpose document edge detection because it requires

25   modification or otherwise specialized design of the hardware for the system. The present

inventors then further considered typical existing scanning devices and came to the

5   realization that many include a cover thereon under which the original document is

positioned. Traditional wisdom suggests that a cover having substantially the same color

as the background of the document contained thereunder, such as a white colored cover

and a white document, is unsatisfactory for determining the edge of the document. For

example, Pasco. et al., U.S. Patent No. 5,818,976, in fact teach the use of a cover that

10   contrasts well with the page. In direct contrast to this traditional wisdom, the present

inventors realized that a cover with the same general color as the document itself may be

used in determining the boundary of a document. In actual systems, the document does

not tend to lay perfectly flush against the cover and accordingly, when the document is

illuminated, a slight shadow is cast by the document onto the cover along a sufficient

15   portion of the edge of the document.

Referring to FIG. 1, a document 10 is supported by a scanning device 12

with a cover 14 covering the document 10, all of which may be flat. Preferably, the cover

is substantially flat. Preferably, a major portion of or all of, the cover has substantially

the same color as the background color of the document to be scanned. More

20   specifically, preferably a major portion of the portion of the cover proximate the edge of

the document preferably has substantially the same color as the background color of the

document to be scanned. The scanning device may be any type of device capable of

obtaining or otherwise sensing an image of the document 10. The document 10 may be

any type of document or otherwise an object that is sensed by the scanning device 12. In

25   addition, the scanner may use a roller or other backing arranged in a manner opposing the

imaging system with respect to the document.

5   Referring also to FIGS. 2A-2C, the document 10 including a portion of the

cover 14 extending beyond the periphery of the document 10 is imaged or otherwise

sensed by the system. FIG. 2A illustrates a scanned document with a skew, together with

horizontal and vertical boundary lines. FIG. 2B illustrates a scanned document with

wrinkles, together with horizontal and vertical boundary lines. FIG. 2C illustrates a small

10   document, together with horizontal and vertical boundary lines. The image acquisition

may be a normal scan, a preview scan at a lower resolution than the normal scan, a

preview scan at a higher resolution than the normal scan or any other type of image

acquisition. The image is normally acquired in a color space that includes red, green, and

blue. Alternatively, any set of one or more colors may be used, black and white, or any

15   other image description scheme. The edges of the document 10 cast a slight shadow onto

the cover, at least a portion of which is likewise sensed. The resulting image together

with the shadow may be processed in any suitable manner to determine the size or

boundaries of the document.

The process described below is suitable for processing image documents

20   in general and particularly suitable for processing the imaged document shown in FIGS.

2A-2C. The process is defined in terms of the document being properly registered with

respect to the top-left corner of the platen. It is to be understood that the process may be

readily extended to the document being located at any position, including a random

position and random orientation. In addition, the number (e.g., one, two, three, four, five,

25   etc) of vertical and horizontal boundary lines (including other orientations of the

boundary lines, such as inclined) may be extended depending on the location and shape

5   of the document. Referring to FIG. 3, after acquisition of the image 20 in a red, blue,

green color space the image 20 is preferably converted from the reg, green, blue color

space to a color space that enhances the luminance of the image at block 40. With a

document that is sensed under relatively uniform illumination, especially when narrow

shadows are to be sensed, it is preferable to process the image therefrom in terms of

10   enhanced luminance. It is to be understood that the image may be processed in any color

space, as desired. The conversion from the red, green, blue color space to luminance Y

may be computed as: $Y=(0.3R +0.59G + 0.11B)$. Preferably the acquired image 20 is

obtained at a lower resolution than the normal resolution used by the system for creating

a copy in order to reduce the memory requirements. The conversion from a triplet color

15   space (e.g., red, green, blue) to a luminance results in a reduction of the data by

approximately a third, which reduces the memory requirements of the system and the

computational complexity.

Preferably after converting the image to a color space that enhances

luminance, the predicted range of values representative of anticipated document boundary

20   edges may be stretched or otherwise enhanced to provide a greater weight, sensitivity, or

otherwise, at block 50. Stretching increases the robustness of the edge detection process

and enhances shadow edges by increasing the differences of pixel values in the range of

likely document edge values and by attenuating edge magnitudes in the color range of the

scanner cover and other data such as text. For example, pixels having a luminance in the

25   range of 190-220 (potential values from 0-256) may be stretched to the range of 170-240

by applying an S-curve. It is to be understood that any modification of the image to

5   enhance image characteristics likely to be characteristic of the edge of a document may

be used. In addition, the image modification by conversion to luminance enhancement,

stretching, if performed at all, may be performed at any time during processing.

Preferably after converting and stretching the image, the image is down

sampled to a lower resolution, such as 75 x 75 dpi, at block 60. Down sampling the

10   image from a 300x150 preview scan resolution to a 75 x 75 resolution results in

approximately an 8 times reduction in the data. This likewise results in a consistent

sampling density for further processing, which provides greater consistency for image

processing and more flexibility in implementing the system on different platforms. A 75

x 75 resolution generally results in no more than 640 x 896 pixels (an A4 U.S. letter-sized

15   scanner platen is assumed without loss of generality). In addition, down sampling the

luminance enhanced data is less computationally intensive than down sampling the

original image data. For example, a 1 x 4 box filter average in the horizontal direction

and a two tap IIR filter in the vertical direction may be used.

The image resulting from the down sampling 60 may subsequently be

20   divided into row strips (e.g., 32-pixels high) and column strips (e.g., 32-pixels wide). For

each row strip, a set of contiguous sub-rows may be selected, such as 8, 16, or 32 rows.

For each column strip, a set of contiguous sub-columns may be selected, such as 8, 16, or

32 columns. In essence, the down sampled image 60 is partitioned into a set horizontal

strips consisting of a group of rows, and into a set of vertical strips consisting of a group

25   of columns. It is to be understood that any number of pixel strips, any number of sub-

strips, contiguous or non-contiguous, may be used. For the illustrated example, 8 element sub-strips are used, simply for ease of illustration.

The transverse average for each horizontal sub-strip is computed and stored in a horizontal stat buffer, at block 70. It is to be understood that any other statistical measure for each horizontal sub-strip may likewise be used, as desired. In the particular example illustrated, each 8-row horizontal sub-row is 640 columns wide. In the illustrated example there are 28 such 640 element row-stat buffers for the image. An exemplary row stat buffer is shown in FIG. 4A for the image shown in FIG. 2A.

Similarly, the transverse average for each vertical sub-strip is computed and stored in a vertical stat buffer, at block 80. It is to be understood that any other statistical measure for each vertical sub-strip may likewise be used, as desired. In the particular example illustrated each 8-column sub-strip is 896 rows high. In this example there are 20 such 896 element column stat buffers. The number and length of the row and column buffers may be selected, as desired. An exemplary column stat buffer is shown in FIG. 4B for the image shown in FIG. 2A.

The use of column and row statistical buffers permits the simulation of a larger convolution kernel which results in more robust processing and likewise reduces the amount of data. Further, the transverse processing reinforces the image detail in a transverse direction which emphasizes the shadow on the edges of a document. In addition, the relatively tall filter relative to the typical height of the text tends to attenuate the text.

At block 90 a localized 1-dimensional difference operator identifies the edges of the image whose magnitude difference is above a selected threshold. For example, points whose measured local difference along the row (or column) stat buffer is greater than 10 may be considered edges. It is to be understood that any one or multi-dimensional operator which identifies edges in an image may likewise be used. The use of an appropriate operator tends to identify those regions of the image that are candidate regions of the shadow cast by the document. Conceptually, this results in another row stat data structure and another column data structure where edges are identified.

With the potential edge regions of the row stat buffer identified, or otherwise potential edge regions of the image, the total number of potential identified edge features for each transverse column are summed together. The total number of edge features for each transverse column is stored in an accumulation row buffer at block 100. An exemplary accumulation row buffer is illustrated in FIG. 5A, with each vertical line representative of a region of 32 pixels. Similarly, with the potential edge regions of the column stat buffer identified, or otherwise potential edge regions of the image, the total number of potential identified edge features for each transverse row are summed. The total number of edge features for each transverse row is stored in an accumulation column buffer at block 110. An exemplary accumulation column buffer is illustrated in FIG. 5B, with each horizontal line representative of a region of 32 pixels. This results in an increased likelihood of accurately determining the appropriate horizontal and vertical positioning of the edges resulting from the cast shadow. Further, another significant

5   reduction of the amount of data is accomplished, e.g., 28 (rows) x 640 (columns) to 1

(row) x 640 (columns).

While transverse accumulation aids in identification of those regions of

potential shadows, however, the potential skew of the document itself tends to spatially

spread the apparent edge. To compensate for the potentially skewed edge, the data in the

10   accumulator is passed through a smoothing function at block 120, such as for example, a

Gaussian filter [1, 2, 1]. In essence, each particular value is adjusted in accordance with

its neighboring values. The effect is to emphasize values in regions having significant

spatially adjacent or proximate values. An exemplary smoothed data set of the row

buffer is shown in FIG. 6A and an exemplary smoothed data set of the column buffer is

15   shown in FIG. 6B. Alternatively, emphasizing values in regions spatially adjacent or

proximate to one another may be undertaken during other processes, such as the

accumulation process.

The boundary of the image, or otherwise the document, may be

determined based upon the largest value in the accumulator or as a result of the

20   smoothing. Another technique to determine the boundary of the image is to select the

outermost value greater than a sufficient threshold at block 130. In addition, the system

may likewise determine the boundary region of the image, text or other content on the

document. However, larger images tend to have larger smoothed accumulator values,

while smaller images tend to have smaller smoothed accumulator values. This difference

25   in the maximum values tends to make it difficult to select an appropriate threshold value.

To overcome the thresholding dilemma, the present system may incorporate a threshold

-11-

5   that is expressed as a percentage (or other statistical measure) of the maximum observed

row or column edge count (or other criteria). This permits a single threshold to be used

for both the horizontal and vertical boundaries, even with different sensitivities in each

direction. The horizontal boundary of the document may then be considered as the right-

most row edge count above the row-scaled threshold. The vertical boundary threshold

10   may then be considered as the bottom-most column edge count above the column-scaled

threshold. Moreover, a single threshold value is likewise generally scale and

directionally invariant.

The terms and expressions which have been employed in the foregoing

specification are used therein as terms of description and not of limitation, and there is no

15   intention, in the use of such terms and expressions, of excluding equivalents of the

features shown and described or portions thereof, it being recognized that the scope of the

invention is defined and limited only by the claims which follow.

20